Center for Independent Experts (CIE) Independent Peer Review Report on STAR Panel 1 Meeting – Pacific Coast Lingcod and Pacific Ocean Perch, June 2017

N. G. Hall

Unit 2
2 Wexford Street
Subiaco
Western Australia 6008
Australia
Phone: +61 8 9401 6891
Email: nghall@iinet.net.au

August 2017

Contents

1. Executive Summary

A Stock Assessment Review (STAR) Panel was convened at the NOAA Fisheries Northwest Fisheries Science Center (NWFSC), Seattle, from June 26-30, 2017, to review draft stock assessments for Pacific coast lingcod (*Ophiodon elongatus*), and Pacific ocean perch (*Sebastes alutus*). The STAR Panel found the assessments to be of a high quality and accepted the final base models proposed by the Stock Assessment Teams (STATs). Insufficient time was available during the review meeting for the production of final decision tables and management advice. Based on the results and diagnostic outputs produced during the course of the meeting, however, I concur with the decision by the STAR Panel that these final base models should produce output that facilitates provision of the best scientific advice that is currently available regarding the status of these two stocks.

It was recognised that, in deciding upon the structure to be used as the final base model for each stock, alternative states of nature have been removed from consideration. Key decisions made during the review of the assessments for the lingcod and Pacific ocean perch fisheries are discussed below.

A key decision for the lingcod assessments of the populations in both the north and the south was to exclude marginal age compositions from the assessment models. Evidence from fisheries data suggested that these were not representative of the age composition of the modeled population. For the south, another key decision was to remove the recreational fishery indices, as depletion estimates at the end of the time series were quite sensitive to these data. This decision acknowledged that the fishery-independent indices were likely to be more reliable indicators of population abundance than the fishery-dependent indices, and that the different fisheries indices varied in their spatial coverage of the southern region. The alternate scenario produced when these indices are used in the model for the south should be included in the final assessment report as a sensitivity run. The decision to only estimate additional standard deviation (SD) for fishery CPUE indices, but not for survey indices was also important as this ensured that signals regarding trends in abundance were not compromised by down-weighting to improve fits to composition data. For the south, the assumption of linear ramping of early commercial catches from zero to the value of the first recorded landings in 1931 provided a more realistic trajectory of spawning biomass in early years of the fishery, but did not affect the status determination.

The key decision for the Pacific ocean perch assessment was whether to retain or remove the Triennial survey data as there was clearly tension between the values of parameter estimates supported by the NWFSC indices and those supported by that data. The decision was made that, for the final base model for the 2017 assessment, it would be appropriate to recognise that the NWFSC survey is considered by the STAT to provide the best representation of the Pacific ocean perch population. The alternative scenario represented by including the Triennial survey data should be presented in the final assessment report as a sensitivity run. It was recognised that the decision made by the STAT to include the conditional ages at length (CAAL) for the NWFSC shelf-slope survey in the final base model would influence the estimates of spawning output produced by that model. As with the decision regarding the Triennial survey data, it would be appropriate to present a sensitivity run in the final assessment report for the case when these CAAL data are excluded and fixed growth parameters are input to the model.

The lack of a completed study that validates the ageing procedure for lingcod is a source of considerable uncertainty. If the ages assigned to individuals of this species are incorrect, the status of the north and south stocks of lingcod will need to be re-assessed. The potential that the movements and spatial distributions of both lingcod and Pacific ocean perch are age-dependent also requires investigation, as this has implications for obtaining random samples of conditional ages at lengths.

The review meeting was well organised, and documents were available on time. The Stock Assessment Teams are commended for their efforts and thanked for their willingness to respond to panel requests.

2. Background

2.1. Overview

A Stock Assessment Review (STAR) Panel was convened at the NOAA Fisheries Northwest Fisheries Science Center (NWFSC), Seattle, from June 26-30, 2017, to review draft stock assessments for Pacific coast lingcod (*Ophiodon elongatus*) and Pacific ocean perch (*Sebastes alutus*). The agenda for the Review Workshop is presented in Annex 3 of Appendix 2.

The Statement of Work provided to Dr Norm Hall by the CIE is attached as Appendix 2. This required that, in addition to satisfying the requirement for STAR Panel members to participate in the review and conduct an independent peer review of each assessment, Review Panel members should assist the Review Chairman in preparing a STAR Summary Report of the review, and each should also prepare an independent CIE report of the assessments and the review process. This CIE review report, which is prepared in accordance with the last of these requirements, describes my evaluation of the assessments and the review process.

Prior to the STAR Panel Meeting, the stock assessment documents and other background documentation had been made available to Panel members. A list of these documents is presented in Appendix 1.

2.2. Terms of Reference

The terms of reference for the stock assessments of the lingcod and Pacific ocean perch are presented in the Statement of Work (Appendix 2), together with the terms of reference for the STAR review of these assessments.

2.3. Panel membership

Details of the Panel Membership and of other key participants for the STAR review of the lingcod and Pacific ocean perch are presented in Appendix 3. In particular, the STAR Panel members comprised:

- Dr David Samson, Chair, Oregon State University/Oregon Department of Fish and Wildlife and SSC member
- Dr Kevin Piner, NOAA Fisheries SWFSC and SSC member
- Dr Yiota Apostolaki, CIE
- Dr Norman Hall, CIE

2.4. Date and place

The STAR Panel convened from 26–30 June, 2017, at the Northwest Fisheries Science Center, Seattle, Washington, to review the 2017 stock assessments for lingcod and Pacific ocean perch.

2.5. Acknowledgments

Thanks are expressed to the various individuals who participated in the review meeting, and who contributed to the stock assessments, for making the review such an interesting and positive experience. The Stock Assessment Teams (STATs) and, in particular, the presenters, M. Haltuch and C. Wetzel, are to be commended for the quality of their stock assessments, and their very competent and professional responses to the Panel's queries and requests. Thanks are also extended to J. DeVore for his input, recording requests to the STATs, and updating the FTP site with presentations, requests, additional documentation, and webinar recordings, to L. Mattes, and L. Zimm for their input, and to S. Miller for ensuring the smooth running of the review meeting.

3. Description of Reviewer's role in review activities

Prior to the review meeting, I downloaded and familiarised myself with the background documentation, and the assessment and draft assessment summary reports for the two species that were the subject of the review (Appendix 1). Subsequently, I attended and actively participated as a Panel member in the review that was held in Seattle, and was rapporteur for the assessment of the southern lingcod population. With other Panel Members, I contributed to the STAR Summary Report, and prepared this report.

4. Summary of findings relevant to the Terms of Reference

*4.1 Pacific coast lingcod (*Ophiodon elongatus*).*

Lingcod ToR 1. Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.

Brief summary of data used for 2017 assessment of lingcod

Although not genetically distinct, the lingcod population off the US west coast is modelled as two stocks, i.e. a northern 'stock' in the waters off Washington and Oregon, and a southern 'stock' off the coast of California. The species is demersal, occupies depths to approximately 450 m and is most abundant in areas of hard bottom with rocky relief, typically occurring at depths less than 200 m. Although there is inshore-offshore movement associated with reproduction, lingcod exhibit high site fidelity and there is limited along-shore movement of older juveniles and adults. Its biological characteristics, i.e. growth and maturity, vary with latitude, offering further evidence of limited north-south connectivity as juveniles and adults.

Lingcod exhibit sexually-dimorphic growth, with females growing to a larger size than males. Growth rates and sizes at maturity vary spatially, with fish at higher latitudes growing more slowly and maturing at larger sizes than those at lower latitudes. Outside the spawning season, females tend to occupy deeper habitat in offshore waters, while males are found in nearshore, rocky reef habitat. For both sexes, size increases with depth. Spawning typically occurs in shallower water (10-40 m) with rocky, high relief habitat. Males move from trawling grounds in late fall, aggregating in areas suitable for spawning. Females apparently move into spawning areas for only a brief period to deposit eggs, with larger, older females spawning first. Males guard the nests, into which the females have deposited their eggs, till the eggs hatch (~5-7 weeks). Larvae are epipelagic, exhibiting vertical diurnal movement, until they reach lengths of ~70 mm, when they settle to soft bottom habitats. By age 2, juveniles have moved from inshore areas into habitats of similar relief and substrate, but lesser depths, as occupied by adults.

An updated functional maturity ogive was used to relate female maturity to length, with fecundity assumed to be proportional to weight. A power curve was used to calculate fish weight from length. Ageing error was determined from double reads by the NWFSC Cooperative Aging Project and Washington State laboratories, but production ageing has still not been validated. Discard mortalities of 7% for fixed gears and 50% for trawl gears were assumed. The distribution of the prior for natural mortality was updated, using collated data from Then *et al.* (2015). Maximum age for female lingcod was assumed to be 21 years, with the prior estimate for female mortality fixed at 0.257 year^{-1}. No prior is available for steepness.

The model for the north considered the reconstructed landings for four fishing fleets, i.e., trawl and fixed gear commercial fleets and Washington and Oregon recreational fleets, where landings reported in WA and OR for these fisheries include fish that may have been caught in either state. Indices of abundance and length and age compositions are available for various periods during the years in which the lingcod in this region have been exploited. These indices were calculated using a spatio-temporal model implemented in the R package, VAST, or a delta-GLM approach. In the south, a Bayesian delta-GLM was also employed to model the presence/absence of lingcod in catches. For the north, available data comprised CPUE data for each commercial fleet (trawl: 1981-1997 (VAST); fixed gear: 2004-2016 (delta-GLM)), and for the Washington (1981-2016 (delta-GLM)) and Oregon (1986-2016 (delta-GLM)) recreational fisheries (reflecting CPUEs from either state). These fishery-dependent indices were complemented by indices of abundance for this stock from three fishery-independent surveys, i.e. Triennial surveys for the early (1980-1992 (VAST)) and later (1995-2004 (VAST)) periods and a NWFSC trawl survey (2003-2016 (VAST)). Discard rates are available for both commercial fleets (both: 2002-2015).

Length compositions for the north are available for trawl landings (1965–1975, 1978-2016) and discards (2004-2015), fixed gear landings (1971, 1978–2016) and discards (2004–2015), catches by recreational fishers in Washington (1979–2016) and Oregon (1980–1989, 1993–2016), the early (1986-1992) and late (1995-2004) Triennial surveys, the NWFSC trawl survey (2003–2016), WDFW research studies (‘Jagielo’: 1996-1997, 2001–2003; indices not used as original data misplaced), and a research study by L. Lam (2016). Marginal age compositions are available from the commercial trawl (1978–2016) and fixed gear (1978–1983, 1986-2016) fleets, the recreational fisheries of Washington (1967–2016), and Oregon (1974–2008, 2012, 2014-2016), the late Triennial survey (1995-2004), and the WDFW research studies (1996-1997, 2001-2003). Conditional ages at length (CAAL) are available from the NWFSC trawl survey (2003–2016) and the research study by L. Lam (2016).

Removals in the south were due to three fishing fleets, i.e., trawl and fixed gear commercial fleets and the Californian recreational fleet. As with the north, indices of abundance and length and age compositions are available in the south for various periods within the history of exploitation. CPUE data are available for the commercial trawl fishery (1981-1997 (VAST)) and both an on-board observer index (early: 1987–1998; late: 2002–2016 (delta-GLM)) and central Californian dockside index (‘Wadsworth index’: 1980-1997) are available for the recreational fishery. Sources of fishery-independent survey data in the south were the early (1980-1992 (VAST)) and late (1995-2004 (VAST)) Triennial surveys, the NWFSC trawl survey (2003-2016 (VAST)), and a NWFSC Hook and Line survey for the southern Californian Bight (‘Harms survey’: 2004-2016 (Bayesian delta-GLM to model presence/absence of lingcod)). Discard rates are available for both commercial fleets (trawl: 2002-2015; fixed: 2002, 2004-2015).

Length compositions for the south are available for trawl landings (1978–1980, 1982–1989, 1993-2016) and discards (2004-2015), fixed gear landings (1978, 1979, 1982, 1983, 1985–1989, 1993-2016) and discards (2004-2015), recreational fishers (1959–1972, 1975–1989, 1987–1998, 1993-2016), the early (1989-1992), and late (1995-2004) Triennial surveys, the NWFSC trawl survey (2003–2016), the NWFSC Hook and Line Survey (2004 – 2016), and from the research study of L. Lam (2016). Age compositions for the south are available for the commercial trawl (1993–1998) and fixed gear (1993, 1994, 1998, 2004) fisheries, and the late Triennial survey (1995-2004). Conditional ages at lengths are available for the NWFSC trawl survey (2003–2016) and the research study by L. Lam (2016).

Brief description of model structure for 2017 assessment of lingcod

The draft base models for the lingcod population in the north and south were implemented, separately and independently, as length-, age- and sex-structured single-area models in Stock Synthesis V3.30.01.05. Although reconstructed landings for the commercial fleets in the south are only available from 1931, for consistency the models for both the north and south run from 1889 to 2016, with forecasts from 2017. An accumulator age of 25 years was used, with data employing a plus group at age 20 years. Length bins in the north ranged from 10 to 130 in 2 cm increments, while those in the south ranged from 4 to 130 cm. Sex ratio at birth was fixed at 1:1. Growth, natural mortality, and selectivities for females and males were sex-specific, with estimates of female spawning biomass providing a proxy for egg production. Growth was modelled as a single growth morph, and assumed to be time invariant. Input sample sizes for length and marginal age compositions for commercial fishery data and surveys were taken as the number of trawl tows or port samples, and number of fish ages was used as the input sample size for the conditional-age-at-length samples and samples from recreational fisheries. Sample sizes were iteratively re-weighted using the Francis (2011) method. The input value of σ_R was iteratively re-weighted to ensure consistency with the empirical estimate. Additional variances were estimated for indices of abundance. Log-likelihoods were estimated for catches, indices, discards, length compositions, age compositions, recruitments, parameter priors, and parameter soft bounds.

Size selectivities, which were estimated using the double normal pattern, were time-blocked for the commercial trawl and Oregon recreational data in the north, and for the commercial trawl and recreational fishery in the south. Retention curves for both the commercial trawl and fixed gear fisheries were time-blocked in both the north and south. Survey catchability estimates were calculated analytically. The documented base models for both the north and south had been corrected by the STAT prior to the STAR, and alternate base models, which the STAT favoured, were also presented for consideration by the STAR Panel.

In the documented base model for the north, natural mortality of females was fixed at the value of the prior for a maximum age of 21 years, i.e., 0.257 year^{-1}, the steepness of the stock-recruitment relationship, h, was fixed at 0.8, and σ_R was fixed at 0.6. In this documented base model, both female length at the maximum reference age and k for female growth had been fixed, but in the alternate base model, only female length at the maximum reference age was fixed. The difficulty encountered in estimating female length at the maximum reference age was possibly due to the paucity of data for older fish in the NWFSC survey and the Lam dataset. Natural mortality of males was estimated, assuming the Hamel (2015) prior. Annual recruitment deviations starting at 1985 had been estimated for the documented base model, but, for the alternative models, recruitment deviations had been estimated from the starting year for the model, i.e. 1989. Lack of information within the data for the south required an assumption for its base model that the values for female natural mortality and steepness would be the same as the corresponding values for the north. For the southern population of lingcod, however, it was possible to estimate all growth parameters.

At the conclusion of the review, the final base model for the north, which was recommended by the STAT and accepted by the STAR Panel, made use of all indices and length compositions, but only CAALs from the NWFSC survey and the Lam study and no marginal age compositions. It was assumed that, prior to the start of the model, the fishery was at an unexploited equilibrium. For this model, h was fixed at 0.7 and M for females fixed at the prior value associated with a maximum age of 21 years, i.e. 0.257 year^{-1}. Female length at maximum age was set to 110 cm and σ_R to 0.55. Recruitment deviations were estimated starting at the beginning of the modeled period, i.e., 1889. Estimates were made of additional standard deviations for fishery CPUE indices but not surveys. Growth parameters, other than that for female length at maximum age, were estimated. Time blocking was maintained, together with the estimation of the natural mortality of males. As time did not allow the STAT to complete all modeling activity during the course of the review, following the meeting, the model was to be re-tuned after fitting.

The final base model for the south, which was recommended by the STAT at the end of the review and accepted by the STAR, included linear ramping from a catch of zero to the catch reported in the first year for which commercial landing data were available. Female natural mortality was fixed at 0.257 $year^{-1}$ and steepness to 0.7, i.e., the values for the north. For this model, σ_R was set to 0.75. The model made use of all fishery-independent indices and included the trawl CPUE index with estimation of additional standard deviation. It retained the length compositions and CAALs from both the NWFSC survey and the Lam study, but removed all marginal age compositions. As time did not allow the STAT to complete all modeling activity during the course of the meeting, following fitting, the model was to be re-tuned and all diagnostic output produced.

For both of these final north and south base models, an attempt was to be made, after the meeting, to fix *h* and estimate female natural mortality, and vice versa.

Lingcod ToR 2. Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.

In addition to updating the landings and survey data available for the 2009 assessment with the data for subsequent years, fisheries data had been reconstructed and extended back to 1889 in the north. These refinements to the accuracy of the landings data employed by the assessment models will have improved the accuracy of the stock assessment. There was concern by the STAR Panel, however, that commercial exploitation in the south commenced earlier than 1931, the year in which landings first became available. Future assessments would be improved if data for this earlier period were to be resurrected. For the 2017 assessment, the assumption of a linear ramping of landings was considered appropriate.

Disaggregation of the commercial fisheries data input to the 2017 into trawl and fixed gear components and splitting of the recreational fleet in the north into Oregon and Washington fleets allowed the different selectivities of these fleets to be taken into account, thereby improving the quality of the assessment. Inclusion of additional survey indices, and length and age compositions within the 2017 assessment model had the potential of providing additional information on the population dynamics of the lingcod stock.

For the 2017 assessment, new indices of abundance had been estimated for the NWFSC and Triennial surveys, and for the commercial CPUE data, using a spatio-temporal delta model, implemented as the R package, VAST. There was concern, however, that, when undertaking these analyses, the data for both the north and south were analysed simultaneously. This opened the possibility that data from one region might influence the estimates for the other. Such a result would be inconsistent with the decision that had been made for the assessment to treat the lingcod populations in the two regions as two separate and independent stocks. It might also explain the greater similarity of the NWFSC survey indices produced by VAST for north and south than that of the indices produced using design-based analyses. It was noted by the STAR that the assessment document was required to include a table comparing the results from VAST with those from design-based analysis.

Spatial patterns of the Pearson residuals for encounter and catch rates for the PacFIN logbook trawl data exhibited marked inter-annual changes, particularly in the south. A possible explanation proposed by the STAT was that, because of the large quantity of data, it was necessary to employ fewer knots in the analysis. Further investigation is warranted, however.

The 2017 assessment document fails to advise that the hard structure used when ageing lingcod is a spine from the lingcod. The document does advise, however, that an ageing validation study has not yet been completed for lingcod and that more work is required to identify potential biases in production ageing of lingcod. If ageing is invalid or biased, doubt is cast on the age composition data and growth curves.

A plot of the weight-length relationships for unsexed lingcod lay below those for both females and males, and that for females lay below the curve for males. Such results are unexpected and require further investigation and explanation.

The STAT advised in its assessment report that many of the compositional data for the commercial and recreational fisheries lack details of the number of fish sampled out of those landed in a given trip, and are used without expansion to the sample level. This introduces bias into the resulting composition data, as it is assumed that the composition that is input to Stock Synthesis is representative of the composition of the catches that were sampled. Consideration should be given to excluding data for such trips when determining length and age compositions.

Reluctance by fishers to allow cutting of fish to extract spines or identify sex was identified during the STAR as a major source of ageing error, resulting in bias in samples of marginal distributions of lingcod ages for fisheries data.

No age compositions are available for the commercial fisheries of the southern stock since 2004.

It is assumed by Stock Synthesis III that length distributions are random with respect to the catches from which the samples are drawn. It is also assumed that conditional ages at length are random samples of the ages of the fish of that length within the population. Factors that affect the randomness of CAALs include age-dependent movement and spatial distributions. The potential for such age-dependent distributions should be explored, e.g. by comparing CAAL distributions among different depths and different months within a year.

The STAT identified the NWFSC survey as producing data that, based on spatial extent, properties of the gear, and timing, are best representative of the lingcod population.

Lingcod ToR 3. Evaluate model assumptions, estimates, and major sources of uncertainty.

Prior to the STAR, the STAT had undertaken a bridging analysis to ensure that change from the earlier to the current version of Stock Synthesis had not had an unexpected impact. It had also undertaken numerous explorations of the sensitivity of the model to various alternative assumptions regarding model structure and data, details of which were listed in the draft assessment document. For the documented base model, this exploration included fixing female natural mortality and h at different input values, removing individual indices, length and age data sets. Model output had proved sensitive to the starting year for calculation of recruitment deviations. Jitter analyses had demonstrated that, while the model for the north appeared robust to alternative initial values of parameter estimates, that in the south was not as reliable with 52 of 100 jittered starting values producing improved log-likelihoods, some of which represented improvements in fit of up to 5 log-likelihood units.

Although alongshore movement of older juveniles and adults is likely to be limited, there is potential for greater movement of larvae and younger juveniles. Given genetic similarity throughout the spatial extent of the distribution of lingcod, some connectivity would be expected between the North and South stocks, and between the stocks off the west coast of the US and those off Mexico and Canada. The assessment models assume no connectivity, but the sensitivity of the assessment to failure of this assumption should be investigated. Sensitivity runs to explore the implications of truncation of the population at the Mexican and Canadian borders, and of ignoring potential contribution to recruitment through connectivity with other regions, should be undertaken in future assessments.

The implications of male guardianship on the stock-recruitment relationship are not considered within the model. It is possible that larger males are more successful than smaller males in protecting the eggs within the nests.

The possibility of temporal or spatial variation in the parameters of growth and length at maturity was discussed. It was noted that use of the latest data when estimating growth and maturity may introduce problems if these processes have changed over time. As age and length data for the north extend back to the late 1970s, there is opportunity to explore whether growth has changed since that time.

The question was posed by the STAR Panel as to whether the catch and associated index of abundance provide any information on trend and scale via the production function, i.e., an assessment employing an age-structured production model where growth is specified externally. Such assessment might determine whether the observed trends were the result of fishing. If this is the case, other processes could then be implemented to match the age and length compositions. If not, it is the recruitment variations that are producing the variation in the indices and information on scale must be drawn from the composition data. In this case, if age and length compositions are available, and ages are considered reliable, then it may be preferable to ignore the length data and make use of catch at age removals in the model. It was suggested that the STAT should be requested to explore the fitting of such an age-structured production model.

Use of both marginal lengths and marginal ages introduces the issue of 'double counting'. The STAT advised that this was dealt with in the 2017 assessment by re-weighting lambda. The STAR Panel advised that another alternative would be to use ages in years where ages are available, and lengths in those years when no ages are available.

The trends in spawning biomass for both the north and south models were sensitive to the starting years for estimation of deviations. An earlier starting year allows the initial age composition to be estimated more readily such that it is consistent with the data in the period when compositional data or variation in indices become available to inform recruitment deviations.

It appeared from the likelihood profiles of the models proposed by the STAT that length and age compositions were pulling estimates to different locations in the parameter space and thus are in conflict. If this is the case and an additional variance is estimated for the indices, the model is likely to be conditioned on the combination of its structural assumptions and the length and age compositions, fitting the indices as best these allow, i.e., imposing the view that the most representative data are the compositional data rather than the indices. The trend in spawning biomass for both the north and the south models is sensitive to female M and h. Likelihood profiles for the alternative model for the north revealed the tension between age and length compositions, while those for the alternative model for the south were more chaotic and less informative for each dataset. The Panel advised that the likelihood profiles would be more informative if broken down by both data source and data type.

Although no convergence issues were identified in the jitter analysis of the documented base model for the north, a considerable fraction of jitter runs of the documented base model for the south resulted in improved log-likelihoods, with the improvement being sometimes up to 5 units. Consideration should be given to the use of different phasing and the jitter analysis repeated when a new base model has been determined.

Analyses undertaken by the STAT in response to discussions with the STAR Panel, and requests subsequently made by the Panel, led the STAT to modify and refine the assessment models for the north and south, resulting in models that, at the conclusion of the review, the STAT proposed as final base models for these two areas.

Likelihood profiles were produced for $\ln(R_0)$, steepness and natural mortality of females, key parameters in determining stock productivity and status. Tension exists between data sets for which the minima occur at considerably different values of $\ln(R_0)$, steepness and natural mortality of females, and the profiles differ by more than 2 likelihood units from the minimum of the profile for the total. In the case of the model for the north, there is insufficient information in the data to estimate steepness. The likelihood profiles for the indices, length compositions and CAALs converge on an estimate of female natural mortality of 0.3 year^{-1}, suggesting that the specified fixed value of steepness requires a female mortality estimate at least as large as this upper bound on natural mortality for females. With both female natural mortality fixed at 0.257 year^{-1} and steepness at 0.7, $\ln(R_0)$ is fairly well determined, but, given the model structure that has been assumed, there is tension between the value favoured by the CAALs and that favoured by both the length compositions and indices. The likelihood profiles for steepness in the south model appear very sensitive to the CAALs of the

Lam data, suggesting that these data would be influential if h was to be estimated. As with the model for the north, for the given model structure and values of the fixed parameter, the value of female natural mortality favoured by the data is at least as great as the upper bound, 0.3 year^{-1}. Likewise, as with the north, $\ln(R_0)$, is fairly well determined, but, given the model structure that has been assumed, there is tension between the value favored by the CAALs and that favored by the indices.

Following examination of the likelihood profiles, the decision was made to employ spawning biomass in 2017 as the key axis of uncertainty for both the north and south. Values of M that predicted the 25th and 75th percentiles of the distribution of spawning biomass in 2017 for the base model were to be determined. Management advice would then be determined using the outputs from the base model and those obtained by setting female M in the base model to these two selected estimates of female M. In the event that the STAT should find that the results of such an approach were infeasible, an alternative approach was identified for possible use. The likelihood profiles for the base model suggested that the profile for $\ln(R_0)$ could be employed as the key axis for uncertainty, and that, using the profile, values for $\ln(R_0)$ representing the 25th and 75th percentiles of uncertainty for this variable could be determined. By fixing $\ln(R_0)$ to each of these two values, then estimating female natural mortality (or steepness), the associated trajectories of spawning biomass or of the ratio of spawning biomass to initial spawning biomass could then be calculated.

Lingcod ToR 4. Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.

There was concern by the STAR Panel that commercial exploitation in the south commenced earlier than 1931, the year in which landings first became available. For the 2017 assessment, the assumption of a linear ramping of landings was considered appropriate, and it was suggested that such ramping should be introduced to the model for the south.

For the 2017 assessment, new indices of abundance had been estimated by VAST (using a spatio-temporal delta model) for the NWFSC and Triennial surveys, and for the commercial CPUE data, by analysing the data for both the north and south simultaneously. The Panel recommended that the VAST analyses should be re-run, treating the data for the north and south separately. The STAT later advised that this resulted in an improvement in speed of analysis and the ability to use a greater number of knots in the analysis, thereby improving spatial resolution.

Reluctance by fishers to allow cutting of fish to extract spines or identify sex introduces bias into samples of marginal distributions of lingcod ages. Such potentially biased age data should be excluded from use in the assessment.

Lingcod ToR 5. Determine whether the science reviewed is considered to be the best scientific information available.

The final agreed base models for the north and south lingcod populations were accepted by the STAR as well structured, had been thoroughly investigated by the STAT, and are considered the best scientific information currently available for the formulation of management advice.

Lingcod ToR 6. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.

Short term

Data for the south prior to 1931 should be reconstructed.

An ageing validation study has yet to be completed for lingcod. If the method is invalid, doubt is cast on the age composition data and growth curves. It is recommended that an ageing validation study should be given high priority and be completed prior to the next benchmark assessment.

A plot of the weight-length relationships for unsexed lingcod lay below those for both females and males, and that for females lay below the curve for males. Such results are unexpected and require further investigation and explanation. It is recommended that such investigation be undertaken prior to the next assessment.

Maps showing the spatial extent of the distributions of lingcod and each of the surveys, and of logbook coverage, should be produced, together with maps identifying areas where different habitats might affect, to different degrees, the effectiveness of surveys or fishing with different gears. The extents of overlaps and possible conflicts among data from different sources should be evaluated and the extents to which the various survey and CPUE indices and compositions are representative of the abundance and compositions, respectively, of the population of lingcod should be assessed. The implications of this exploration for the data that are used in the assessment models, and the treatments by the models of those data, should be discussed.

The STAT advised in its assessment report that many of the compositional data for the commercial and recreational fisheries lack details of the number of fish sampled out of those landed in a given trip, and are used without expansion to the sample level. Consideration should be given to excluding data for such trips when determining length and age compositions. It is recommended that, prior to the next benchmark assessment, compositional data should be reconstructed by excluding fish that were sampled in trips for which there are no recorded details of the number of fish sampled out of those landed in a given trip.

It is assumed that the marginal compositions that are input to Stock Synthesis are representative of the compositions of the catches or populations that were sampled. Discussions during the review suggest that fisheries composition data may not always be randomly sampled, or expanded from trip level appropriately. The approaches used to collect samples, and expand samples to whole of catch or survey level, should be reviewed, and compositional data subjected to statistical analyses to assess whether there are any indications of bias in sampling.

A study of existing survey and fisheries data should be undertaken to assess whether, for lingcod of either sex or combined sexes (for samples where it is not possible to cut fish to determine the sexes of the fish), conditional age distributions at different lengths vary spatially or temporally, thereby exploring the assumption that, in conditional age-at-length data, ages of sampled fish at a given length are random samples of the distribution of ages at that length in the population. If non-random, construction of marginal length compositions should account appropriately for the non-uniform sampling of age by length.

Consideration should be given to whether expansion of composition samples for surveys and commercial catches should be modified from a design-based approach to one that accounts for the spatio-temporal distribution of fish determined using the VAST approach.

Concern was expressed during the review that methods employed when calculating the Wadsworth indices might not be readily reproducible. If indices calculated from the central Californian dockside data are to be used in future assessments, the data employed and methods used to calculate the indices should be re-examined and analyses reworked such that the approaches employed may be critically assessed and refined.

Longer term

No age compositions are available for the commercial fisheries of the southern stock since 2004. A research study to obtain such data should be undertaken.

Extend the stock-recruitment relationship for lingcod such that it accounts for the abundance and size distribution of the mature males, which, through their guardianship of nests, ensure survival of eggs.

Undertake further study of the female reproductive process, exploring, in particular, whether females produce multiple batches of eggs, and how fecundity of mature females is related to their body mass.

A model that allows exploration of the implications of connectivity between the north and south populations of lingcod, and between the US and Mexican or Canadian populations, should be implemented. If not already existing, collaborative arrangements with fishery scientists from Mexico and Canada should be established to facilitate collation of the lingcod data from these nations, which would be required for input to such a model.

Lingcod ToR 7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

At the request of the STAR Panel, the STAT produced a table providing details of time blocks used for selectivity and retention, information that was lacking in the assessment report.

To check whether growth had varied over time, the STAT provided box and whisker plots of the NWFSC survey length-at-age data from early and late in the time series. For female lingcod in the north, the rate of increase in length over the first four age classes appeared slightly greater in the later than earlier period but uncertainty bounds for the two periods overlap. There appear to be more older females in the more recent period, with lengths at age appearing to have approached an asymptote. In the earlier period, lengths at age had not approached their asymptote as closely as in the later period. The plotted data for the earlier period support the decision by the STAT to fix the length at maximum age when growth had not slowed sufficiently at older ages to facilitate estimation of this growth parameter. The patterns were similar in the south, with more older fish in the more recent than earlier period. For both the northern and southern stocks, there was no strong evidence of a large change in growth between the two periods. The possibility that growth of the northern and southern lingcod may have changed over the period considered by the assessment models cannot be discounted, however.

The STAT was requested to run the models with h fixed at 0.7 and female M fixed at the prior value for a maximum age of 21 years. A steepness of 0.7 is used for similar species (e.g., cabezon). When the assessment model was refitted to data for the northern stock of lingcod, with fixed h=0.7, M for females fixed at the prior value associated with a maximum age of 21 years, and fixed growth, spawning biomass in 2017 was estimated to be depleted to ~45%. For the southern stock, when the same values were employed for steepness and natural mortality of females and growth was fixed, depletion in 2017 was estimated as ~60%. The Triennial survey index for 2004 was poorly estimated.

The STAT was requested to create combined sex length and age composition data for all commercial fishery samples in the northern model from early years through 1991 and to confirm that length compositions of aged fish are representative of unaged fish. The rationale for this request was that there is an assumed sex ratio for these samples, which may be in error if samples of sexed fish do not include all fish. Age compositions should be representative of fishery selectivity. The trend in the estimates of the spawning biomass of the northern stock that resulted from input of pre-1992 comp. data for PacFIN as unsexed rather than sexed data differed only slightly from that for the model with h=0.7 and M=0.257 year^{-1}. Histograms of aggregated age and length compositions of lingcod collected from the catches taken from the waters from Washington and Oregon by commercial fishers strongly suggested that age compositions were unlikely to represent ages of fish randomly sampled from the fish sampled to produce length compositions. The STAR Panel suggested that this might be indicative of an age-dependent movement or distribution of lingcod.

The STAT's initial examination of the histograms of fish in the aggregated (over time) length and age samples of commercially-caught lingcod from the northern stock appeared to support the view that, taking growth into account, there was agreement between the numbers, i.e. age at length samples were representative of population age at length. Following discussion, however, it was agreed that pinpointing the source of the apparent bias in the composition data would require more detailed examination of data for each year (results of such detailed examination are described below). A suggestion that the age composition data should be excluded from input to the assessment model was rejected as both the STAR Panel and Assessment Team were reluctant to lose valuable information potentially available from any unbiased age composition samples from the commercial fishery. The STAR Panel endorsed further investigation to identify and exclude all commercial fishery age composition data likely to be unrepresentative of the ages of the fish in the length composition samples from which they were drawn.

PacFIN data from Washington and from Oregon within each year for numbers of lingcod in length composition data in different length classes and numbers of fish in those length classes for which ages were available were collated and compared. It is assumed that, if correctly expanded to catches and appropriately combined, length compositions are representative of the length composition of the stock. Confirmation should be sought that, regardless of whether or not catches were recorded as dressed, the approaches applied when expanding the samples were correct. There was convincing evidence in the PacFIN data that, following 2010, the proportions of small fish in length samples from Washington that had been aged exceeded the proportions of larger fish that had been aged. The proportions of aged fish in the different length classes in the annual length samples from Oregon were highly variable, but again it was clear that age samples were not consistently representative of the length samples from which they had been drawn. If samples from either state were affected, the combined data for the northern stock would be biased. The STAR Panel concluded that age compositions were often not representative of the length compositions, and thus associated marginal age compositions for this stock are biased and do not represent the age composition of the stock. Subsequent discussion revealed that permission to cut sampled fish, particularly larger fish, to extract spines for ageing and to sex those fish was often refused, resulting in biased samples. It is likely that similar difficulty was encountered in obtaining samples for ageing from the southern stock, and thus marginal age compositions for this stock were also likely to be biased. Indeed, an instruction had been given in the south, that, when drawing age samples, small and large fish should be targeted, i.e. proportionally oversampled. If such bias is present in any annual data, use of marginal age composition from the affected year is likely to introduce tension among the different data sources in the stock assessments for lingcod. More detailed examination of the age and length composition is necessary to determine whether the bias extends to the age data for all years.

If age-dependent movements occurred, or the spatial distribution of the stock is age-dependent, conditional age at length (CAAL) samples collected from surveys or catches are also likely to be non-representative of the overall conditional distribution of ages at length. In future research, consideration should be given to comparing CAAL distributions from different depths or regions to explore whether these are consistent with the hypothesis that CAAL data are representative of the overall CAAL distribution for the stock.

The STAR Panel advised that, for the current assessment, CAAL data but not marginal age compositions should be used.

Given there are large differences in attributes of the northern and southern models (modeled as separate stocks), it makes sense to separate all data sources. Accordingly, the STAT was requested to run separate VAST runs for the north and south data for all surveys (NWFSC and Triennial) in the northern and southern models, and compare the results from the separate and combined analyses graphically. These would become the datasets used in the final base models. Following this analysis, the STAT advised that, by separating the analyses, the VAST analysis could be run at higher spatial resolution. Although scales differed, there was considerable overall similarity in the trends of the estimates of the survey indices for each region. The slight difference in the trends for the early and late Triennial survey indices indicated, however, that the extent of similarity between the trends of the results from VAST analyses of combined and spatially separated survey data depends on the nature of those data. In the case of the Triennial survey, due to the nature of the bottom, only ~50% of intended sample locations can be trawled safely. The STAR recommended that, for consistency with the decision to model lingcod as two separate stocks, VAST analyses should analyse data for each stock separately. Use of survey indices calculated separately produced a slight change in scale of spawning biomass and only a slight change in the trend. The STAR Panel concluded that it was appropriate to use the VAST indices calculated separately for the northern and southern stocks rather than in a single analysis of data from both stocks.

As the fixed gear fleet currently has composition data for the same time period, the STAT was requested to drop the WA Research comp. data. Removal of these data produced only a slight change in the estimates of spawning biomass. The STAR Panel endorsed this modification to the model for the northern stock.

The marked jump in catch in 1931 in the south (i.e. apparent abrupt start of the fishery) is implausible. Accordingly, for the model for the south, the STAT was requested, as a sensitivity run, to introduce a linear ramp of catch from a value of zero (in the 1st year of catch in the northern model) to 1931 when catches are documented. Assuming a linear ramping of catch from zero to the first year for which landings from the southern stock are available produced a more realistic decline in spawning stock in the very early years of the modelled time period. The STAR Panel advised that it would be appropriate to include such linear ramping in the final base model for the southern stock of lingcod and that, for future assessments, reconstructed catches should be extended to earlier years.

As there may be information to inform recruitment earlier in the model, the STAT was asked to provide a model run with recruitment deviations estimated from the beginning of the model. The STAT advised that, following the modifications to the models for the northern and southern stocks, the difficulty of estimating recruitment deviations from the starting year of the model had been resolved. For the northern stock, such estimation changed the trajectory of estimates of spawning biomass, with greater variation between 1950 and 1980, and a 2017 depletion of ~36%. The bias adjustment ramp for recruitment deviations was examined and appeared appropriate. The fit to the NWFSC survey indices was improved and the fit to the early Triennial survey indices looked good. The Triennial survey index for 2004 was still poorly estimated. In the case of the southern stock, however, the estimate of initial spawning biomass of the southern stock was markedly reduced when recruitment deviations were estimated to the starting year of the model. Spawning biomass of this stock remained lower than the estimates produced by the version of the model employed when combining the sexes for commercial fishery samples until ~1980, declining to slightly lower levels than the estimates of that earlier model, before recovering to a reduced extent to a 2017 depletion of ~30%. Such response by the estimates of spawning biomass of both stocks, and particularly that of the southern stock, suggests that the input data contain information on age structure that are described better by allowing for recruitment variation in the earlier years of the model. The STAR Panel concurred with the STAT that, for both the northern and southern stocks of lingcod, recruitment deviations should continue to be estimated from the first year.

The STAT was requested to run the north and south models removing all indices except the NWFSC survey. Following this, the STAT was asked to add the triennial survey without the 2004 data point. The purpose of this request was to understand the impact of individual survey series starting with the NWFSC survey, which was considered by the STAT to be the most reliable. The second part of the request was intended to better understand the effect of the Triennial survey with the 2004 data point, which doesn't fit well and may be influencing the overall model results. Removal of all indices except that for the NWFSC survey resulted in a slight reduction in the estimate of initial spawning biomass of lingcod in the northern stock, a decline to a similar level in the 1990s as estimated for earlier model runs, but a slightly greater recovery to a 2017 depletion of ~50%. The STAT advised that compositions were still employed when fitting this model, and had not been removed. It was noted that, although survey or fishery indices may have been excluded when fitting the assessment model, it was useful to examine the extent to which estimates of these indices matched input values. Subsequent addition of the indices for the Triennial survey to the model, without the 2004 value, produced spawning biomass estimates that were indistinguishable from those produced using the NWFSC survey indices. For the northern stock, removal of the option for the model to estimate an additional SD for the survey indices made little change to the trend of the spawning biomass.

For the southern stock of lingcod, removal of all indices other than those from the NWFSC survey resulted in a very marked increase in initial spawning biomass, with a decline by ~1990 to levels similar to those produced other model runs for earlier STAR requests, before a marked recovery to a 2017 depletion of ~80%. Subsequent addition of the indices for the Triennial survey, without the 2004 value, produced a slight reduction in the levels of spawning biomass for all years of the modelled period and in the value of 2017 depletion. Removal of the option to include an additional SD in the survey indices for the southern stock produced a marked change in the levels of the estimates of spawning biomass. When indices were penalized by not adding extra SDs, estimates of spawning biomass reverted to values similar to those obtained for the version of the model fitted when estimating recruitment deviations from the first year of the modeled period, and well below the values produced when dropping all indices other than the NWFSC and Triennial indices. For this model, spawning biomass recovered by 2017 to a depletion of ~30%. The STAR concluded that such evidence of tension among the different indices for the southern stock required further investigation.

The STAT was requested to identify and explore the structure for a possible base model for the north. Following up from the model used in the previous run, a stripped-down model was to be set up, using only NWFSC and Triennial surveys with associated CAAL and length comps, and fishery length comps. Recruitment deviations were to be estimated, starting at the beginning of the modeled period, and no extra SD was to be estimated for surveys. All growth parameters except female length at the upper reference age were to be estimated. Time blocking was to be maintained. Male *M* was to be estimated, and the model was not to be re-weighted. After setting up this model, the STAT was requested to sequentially add the following inputs and, at each step, to compare likelihoods. The inputs were the trawl CPUE index, OR Nearshore CPUE index, WA Rec. Dockside CPUE index, and the OR Recreational Dockside CPUE index. Following this, the STAT was asked to add CAALs sequentially, as follows, for as many time periods as seemed suitable and as time permitted (comparing fits to the NWFSC surveys at each step). The CAALs were those of the trawl compositions, fixed gear compositions, OR recreational compositions (lower priority), and WA recreational compositions (lower priority). The rationale for this request was that an earlier run had provided evidence that the fishery age compositions were not representative of the length compositions. The marginal age compositions are judged by the STAT to be biased. The sequential steps of this request were intended to determine which data were most informative and where data conflicts arise. The hope was that the model, as constructed sequentially, would become the final base and that the results of the requested model runs would determine which data should be included in the model.

The STAT presented a summary of the results for the requested steps, noting that, when fitting the models, no attempt was made to estimate the Hessian. First, likelihoods and parameter estimates were determined for the model based on the NWFSC and Triennial survey indices. Addition of the trawl CPUE index improved the fit to the early and late Triennial survey indices by 11.6 and 6 units, respectively. The likelihood associated with the NWFSC survey index, however, was reduced by over 22 points. Addition of the OR Nearshore CPUE index, the WA Rec. Dockside CPUE index, and the OR Rec. Dockside CPUE index further degraded the fit to the NWFSC survey index with each additional index (by <2 likelihood points at each step). The quality of the fit to the late Triennial survey indices remained relatively unchanged while that of the early Triennial survey indices was first degraded then successively improved as the subsequent two fisheries indices were added.

Composition data from the different sources were then introduced. An attempt to fit the model after addition of the CAAL data for a subset of the annual trawl data failed. Addition of the full set of CAAL data for the trawl fishery improved the fit to the NWFSC survey indices but produced a poorer fit to both the early and late Triennial survey indices. The model traded off the fit to the Triennial survey indices with improved fit to composition data, but the process responsible for this is unclear. Growth parameter estimates, however, were also affected by the addition of the composition data, with, for example, a marked reduction in the estimate of female length at minimum age and in the parameter *k* of the von Bertalanffy growth curve. When fitting only the survey and fishery indices, the only conditional ages at length included were those for the NWFSC survey indices. The SDs for growth increased markedly to accommodate the tension in the data. It is possible that the trade off with the Triennial survey relates to a change in growth between the early and late periods, or a change in selectivity. The attempt to then fit the CAAL data for fixed gear compositions failed, but addition of the CAAL data for OR Rec. compositions and WA Rec. compositions succeeded. The fit to the NWFSC survey indices remained relatively unchanged from the previous model run, while the fit of both the early and late Triennial survey indices first improved then deteriorated. Female growth parameter estimates remained at levels considerably lower than the estimates obtained earlier by fitting the fishery indices.

After examining the likelihoods and parameter estimates resulting from the above exploration, the STAT explored two further runs. For the first, a model employing all survey and fishery indices and only the OR Rec. conditional ages was fitted. The second run used all survey and fishery indices and both the OR Rec. and WA rec. conditional ages. The fit to both the NWFSC and the early Triennial survey indices improved from that produced using a model with only survey and fishery indices. There was relatively little further change when the second set of CAAL data was added.

The STAT provided plots of spawning biomass and spawning biomass relative to unfished spawning biomass to allow comparison of the effect of the various scenarios on parameter estimates. Inclusion of the fishery indices reduced the extent to which the northern stock of lingcod was depleted in the 1990s and raised the level to which, in 2017, the stock was depleted from ~30 to ~40%. This change in trajectory appears to explain, at least in part, the marked change in likelihood when the first of the fishery indices was added to the model. Addition of the CAAL data to the model produced spawning biomass trajectories that differed markedly from those of the survey and fishery indices.

The inconsistencies between the survey and fishery indices and the information in the CAAL data that are evident in these tables of results and in the trajectories of spawning biomass could be due to factors such as changes in the growth of individuals in earlier and later periods, changes in selectivity, the appearance of larger than average recruitment at the end of the period, or unrepresentative samples of ages at length for the population as a whole. The latter might arise from age-dependent ontogenetic movements or spatial distributions. The STAR Panel endorsed the STAT decision, based on the results of the analyses they had undertaken to respond to this STAR Panel request, to explore, as a candidate base for northern lingcod, a model employing all survey and fishery indices and the CAAL data for the survey indices but not fisheries.

As with the request to explore the structure for a possible base model for the north, the STAT was requested to undertake a similar process, following the same steps as above, to explore the structure for a possible base model in the south. A linear ramp-up of catches from the starting year of the model was to be maintained. As with the results for the northern stock, but to a lesser extent, the likelihood of the NWFSC survey indices for the southern stock was reduced as the fishery indices were successively added. While the fit of the late Triennial survey indices similarly deteriorated with successive addition of fisheries indices, the fit of the early Triennial survey indices gradually improved as the trawl and recreational observer fishery indices were added. During the discussion, it was discovered that the recreational hook and line survey data had incorrectly been specified to SS3 as biomass rather than number; this error would be corrected for subsequent runs. It was noted that spatial coverage of the fisheries indices differed. Addition of the CAAL data resulted in a marked deterioration in the fit of the NWFSC survey indices, and, to a lesser extent, deterioration in the fit of both the early and late Triennial survey indices. As with the northern stock, the value of the coefficient, k, of the von Bertalanffy growth curve for females declined when the CAAL data were included in the model stock.

Discussion of the results revealed that, although the Wadsworth indices had been employed in previous assessments, methods employed when calculating the indices might not be readily reproducible. For the current assessment, with the information and data available, it would be difficult to defend the use of these indices when calculating the likelihood and fitting the model. If the Wadsworth indices are to be used in future assessments, the data employed and methods used should be re-examined and analyses reworked such that the approaches employed may be critically assessed and, if necessary, refined.

Examination of the trends in the time series of estimates of spawning biomass, and ratio of spawning biomass to unfished spawning biomass demonstrated that, for the southern stock, successive addition of the fisheries indices reduced the extent to which the stock became depleted in the late 1990s and decreased the value of 2017 depletion from ~60% when the assessment model employs only survey indices to ~24%, i.e., slightly less than the minimum stock size threshold, following addition of all fishery indices

Inclusion of CAAL data resulted in markedly different trajectories of spawning stock biomass and ratio of spawning biomass to unfished spawning biomass, demonstrating the inconsistency between the survey and fisheries indices and the CAAL data. As with the northern stock, such inconsistency could be due to factors such as changes in the growth of individuals in earlier and later periods, changes in selectivity, the appearance of larger than average recruitment at the end of the period, or unrepresentative samples of ages at length for the population as a whole. The latter might arise from age-dependent ontogenetic movements or spatial distributions.

Following examination and discussion of these results, the STAT recommended removal of the recreational fisheries indices from the model that would be developed to serve as the final base model for the south, with the alternate version of reality represented in the final assessment report by a sensitivity run including these fishery-dependent indices. This decision would recognise that the fishery-independent data were more reliable than the fishery-dependent data, and that the fishery-dependent sources of data varied in their coverage of the area occupied by the southern population of lingcod.

The STAR Panel endorsed the decision by the STAT, based on the results of the analyses they had undertaken to respond to the various requests, to explore, as a candidate base model for southern lingcod, a model employing all survey and fishery indices in combination with the CAAL data for the survey indices but not those for fisheries. Thus, a request was made for the STAT, based on the model used in the previous explorations, to explore a potential base model for the north with the following specifications. The model was to include all indices, to retain length compositions, and to include only CAALs from the NWFSC survey and the Lam study. It was to have no initial exploitation, estimate female length at the upper reference age, and estimate extra SD on fishery CPUE indices. The model was to be re-tuned, and full diagnostics were to be produced. If time permitted, the STAT was requested to provide likelihood profiles on fixed female M and fixed h as in original base and, again if time allowed, fix h and estimate M and vice versa. The purpose of this request was for the specifications to converge on a consensus base model, for which this request was required as a final check.

Following further exploration, the STAT found that, when fitting all growth parameters, the estimated value of female length at maximum age was reduced from the value that had previously been fixed. At the same time, selectivity for the trawl fishery became asymptotic rather than dome shaped, a result inconsistent with knowledge of this fishery. The STAT decided that it was necessary to fix female length at maximum age rather than leave it free. Accepting that the model was favouring a lower value of this parameter, the STAT set it at 110 rather than 112 cm, the value that had previously been employed and selectivity of the trawl fishery again became dome shaped. The STAR Panel accepted this change. The output produced by R4SS for the model was examined by the STAR Panel, which accepted the recommendation by the STAT that the model should be accepted as the final base model for the north for the 2017 assessment of lingcod.

Likelihood profiles for h, M and R_0 were produced for the model and were used to identify options for specifying the required 12.5 and 87.5% levels of uncertainty. Although the various data sources now in the model carried no information that would facilitate estimation of h or M, the likelihood profile for R_0 suggested that it might be of use in determining values to bracket the uncertainty.

As with the above request to explore a potential base model for the north, the STAT was requested to explore a potential base model for the south. Such a model would build on the structure of the model that had been developed for this area when responding to earlier requests. It would include all fishery-independent indices, include trawl CPUE index with extra SD estimated, include only CAALs only from NWFSC survey and LAM study, and retain length compositions. The STAT was requested to re-tune the resulting model and provide full diagnostics. If time allowed, the STAT was requested to provide likelihood profiles on fixed female M and fixed h, as in original base, and, again if time allowed, to fix h and estimate M and vice versa. These specifications were intended to converge on a consensus base model and this request was required as a final check.

The output produced by R4SS for the specified model was examined by the STAR Panel, which accepted the recommendation by the STAT that it should be accepted as the final base model for the south for the 2017 assessment of lingcod.

Likelihood profiles for h, M and $\ln(R_0)$ were produced for the model to determine options for specifying the required 12.5 and 87.5% levels of uncertainty. The profile for h suggested that it would be possible to estimate h, however, when this was done, the spawning biomass was driven to unrealistic, very low values of depletion. The key variables driving this profile were the Californian recreational and Lam surveys, where, possibly as a result of sampling in areas closed to fishing, the CAALs of the latter survey appeared to have oversampled the older fish. The likelihood profile for M demonstrated that, if an attempt was made to estimate M, the age and length compositions would drive the estimate towards an upper bound.

The STAT presented an updated version of the likelihood profile of $\ln(R_0)$ for the South base model, in which the range of values of $\ln(R_0)$ had been extended. This clearly demonstrated that, for larger values of $\ln(R_0)$, total likelihood continued to increase. The STAT proposed that $\ln(R_0)$ should be used as the axis of uncertainty for both the North and South base models, and to use the likelihood profile to locate low and high values of $\ln(R_0)$ corresponding to the 12.5 and 87.5 percentiles based on the distribution of values of negative log-likelihood centered on the minimum value of negative log-likelihood. Twice the difference between a value of negative log-likelihood for the total log-likelihood and the negative log-likelihood at the minimum would be expected to have a chi-square distribution with 1 df. Thus, using the profile for $\ln(R_0)$ and this distribution, values representing the 12.5 and 87.5 percentiles of the distribution of values of $\ln(R_0)$ could be determined. The base models would then be re-fitted, after fixing recruitment, to estimate the values required for the decision table of the assessment report. The STAR endorsed this approach, but requested that the STAT confirm that the resulting estimates of spawning biomass associated with the 12.5 and 87.5 percentiles of $\ln(R_0)$ encompass the 12.5 and 87.5 percentiles of the normal error distribution around the 2017 estimate of spawning biomass.

Note: The Chair of the STAR Panel subsequently identified that, unfortunately, the STAR Panel had erroneously advised the STAT during the review that the values of $\ln(R_0)$ at the points of intersection of the likelihood profile for the total and a horizontal line 1.15 log-likelihood units greater than the minimum of that likelihood profile would be used as the 12.5 and 87.5 percentiles of the distribution of values of $\ln(R_0)$. The correct difference between the negative log-likelihood at the minimum and the negative log-likelihoods associated with the 12.5 and 87.5 percentiles is 0.66 (Appendix 4).

The final request to the STAT was for it to build decision tables for the north and south based on $\ln(R_0)$ profiles, and choosing the states of nature from the values of $\ln(R_0)$ at the minimum and at the 12.5 and 87.5 percentiles. The goal was to achieve bounds at least as wide as the 12.5 and 87.5 percentiles of 2017 spawning biomass. The rationale for this request was that use of other likelihood profiles to develop decision tables did not provide enough contrast.

Time did not allow the STAR Panel to examine the results of this final request. The error involved in not providing the correct difference, i.e. 0.66, between the negative log likelihood at the minimum and the negative log-likelihood at the parameter values associated with the 12.5 and 87.5 percentiles may not have affected the resultant decision table as the STAT advised in a subsequent email message that it had apparently employed a different approach. For the north model, and presumably also the south, values of spawning biomass at the 12.5 and 87.5 percentiles appear to have been determined using the asymptotic SE. The values of $\ln(R_0)$ that corresponded to these spawning biomasses were then determined using the likelihood profile. Note, however, that details of the calculation and subsequent development of the decision table have not been reviewed. The STAR Panel's error regarding the calculation of the difference between the negative log-likelihood at the minimum and the negative log-likelihoods associated with the 12.5 and 87.5 percentiles should be brought to the attention of the lingcod and Pacific ocean perch STATs, such that they might confirm that their calculations of the decision table results are not affected.

*4.2 Pacific ocean perch (*Sebastes alutus*)*

Pacific ocean perch ToR 1. Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.

Brief summary of data used for 2017 assessment of Pacific ocean perch

The distribution of Pacific ocean perch (*Sebastes alutus*) extends from Alaska to Japan, the Bering Sea and south to Baja California. The stock assessment, which is the subject of this review, is focused on the portion of the population of Pacific ocean perch occupying waters along the west coast of the U.S. from northern California to the Canadian border. There is no evidence of genetic differences within this region, and the possibility of trans-boundary effects of fishing in Canadian waters on the population in U.S. waters has not been ignored. The Stock Assessment Team recognised this possibility and, in their assessment document, reported the result of a sensitivity analysis that assessed the effect of inclusion of Canadian data.

Pacific ocean perch typically occupy waters with depths of 200 to 400 m in summer months, with only the smallest individuals in depths < 200 m and only larger individuals in depths > 300 m. They are rarely encountered south of 40°N. While there is an inshore - offshore movement of females, which is associated with reproduction, there is no evidence of long-shore movement. The STAT advised that ontogenetic movement is probably related to age rather than length. Growth is sexually dimorphic, with females attaining a greater length than males.

The ages of all Pacific ocean perch used in the 2017 assessment were determined from otoliths using the break-and-burn method. Estimated ages of a number of fish (from all west coast sources) ranged from 90 to 102 years. Allowing for error in age estimates, the maximum age of Pacific ocean perch used in the documented base model was assumed to be 100 years. A one-parameter model relating natural mortality to maximum age assuming a log-log relationship was fitted to the data collated by Then *et al.* (2015). The distribution of the prior for natural mortality for a fish with a maximum age of 100 years was updated, using this relationship, producing a point estimate and median for the prior of 0.054 $year^{-1}$. Based on available data, a sex ratio at birth of 1:1 was assumed. An updated functional maturity ogive was used to relate female maturity to length. An updated fecundity-at-length relationship was also employed in the 2017 assessment. Growth curves fitted separately to the lengths at ages of females and males provided the starting values for the growth parameters used within the base model. A power curve was used to calculate fish weight from length. Ageing error was determined from double reads by the NWFSC Cooperative Aging Project with a nonlinear standard error relationship with age reflecting increasing variability of age determinations for older fish. A discard mortality of 100% was assumed. The current prior for steepness h of rockfish is a beta distribution with μ=0.72 and σ=0.15.

Historical landings of Pacific ocean perch from Washington, Oregon, and California, which had been were reconstructed and refined, together with data from PacFIN, data from the At-Sea Hake Fishery observer program, and estimates of foreign catches, were collated and updated to produce revised time series of landings for the commercial fishery (1918-2016), the At-Sea Hake fishery (1975-2016), and the foreign fishery (1966-1976). Removals by surveys were also taken into account in the assessment model. A number of indices of abundance and length and age compositions are available for various periods during the years in which Pacific ocean perch in waters off the US west coast have been exploited. Estimates of CPUE (1956-1973) for the US fleet, which had been calculated by Gunderson (1977), were employed as a fishery-dependent index of biomass. Biomass indices were calculated from the NWFSC shelf slope survey (2003-2016), the AFSC slope survey (1996-1997, 1999-2001), the Triennial Shelf Survey (between 1980 and 2004), and the Pacific ocean perch (POP) Survey (1979, 1985) using the spatio-temporal VAST delta-GLMM model, and from the NWFSC slope survey (1999-2002) using a Bayesian delta-GLMM analysis. Discard rates are available for the commercial fishery for 1985-1987 and for 2002-2015. Discards by the At-Sea Hake fishery are included in the time series of catches for this fishery. It is assumed that no fish were discarded by the foreign fleet.

Biological (length and age) data for Pacific ocean perch for trips by commercial fishers in which this species was caught were extracted from PacFIN, expanded to the state level then combined to fishery level, with input sample sizes calculated from the number of trips and number of sampled fish using the 'Stewart Method'. Lengths and ages for the At-Sea Hake fishery were obtained from the observer program. Length compositions were available from the commercial fishery for 1966-1991 and 1994-2016, and from the At-Sea Hake fishery for 2003-2016. Length compositions for discards were available for 1986, and for 2004-2015. Length compositions were calculated by expansion from samples collected from the NWFSC shelf-slope 'combo' survey (2003-2016), NWFSC slope survey (2001-2002), AFSC slope survey (1996-1997, 1999-2001), Triennial survey (1980, 1983, 1986, 1989, 1992, 1995, 1998, 2001, 2004), and Pacific ocean perch (POP) survey (1979, 1985). At least in the cases of the NWFSCcombo, NWFSCSlope, Triennial, and POP surveys, and presumably also the AFSCSlope, expansion was based upon the stratification of the survey design. Effective input sample sizes were calculated from the number of tows, using the method described by Stewart and Hamel (2014).

Marginal age compositions were available for the commercial fishery (1981-1988, 1994, 1999-2016), the At-Sea Hake fishery (2003, 2006, 2007, 2014), and the NWFSCSlope (2001-2002), Triennial (1989, 1992, 1995, 1998, 2001, 2004), and POP (1985) surveys. Effective input sample sizes were calculated from the number of tows, using the method described by Stewart and Hamel (2014). Conditional age at length (CAAL) data were available, by sex, for the NWFSC shelf-slope survey (2003-2016). Effective input sample size for these data was set equal to the number of fish at each length by sex and year. In subsequent analysis, it was found that the 2016 NWFSC shelf-slope survey differed from previous years for both females and males, with sampled fish being larger at age. Accordingly, the input sample size for this year was reduced to 50% of the number of fish in each length-age bin.

Brief summary of model structure for 2017 assessment of Pacific ocean perch

The model for the Pacific ocean perch population on the US west coast was implemented as a length-, age- and sex-structured single-area model in Stock Synthesis 3.30.03.05. The model employed separate growth and mortality parameters for each sex, representing the dynamics of the stock between 1918 and 2017, with forecasts beyond 2017. Three fishing fleets and five fishery-independent surveys removed fish from the stock. CPUE data from one fleet were available for a short period, while the surveys produced indices of biomass for periods ranging in duration from 2 to 14 years. Length compositions were available for two fisheries (49 and 14 years, respectively), together with marginal age compositions (27 and 4 years, respectively). For the five surveys, length compositions were available for periods ranging from 2 to 5 years, and, for four of those surveys, marginal age compositions were available for periods ranging from 1 to 6 years. For the fifth survey, conditional age at length data for 14 years were available for each sex.

An accumulator age of 60 years was used, with data employing a plus group at age 40 years. Modeled length bins ranged from 5 to 50 in 1 cm increments, with length data ranging from 11 to 47 cm. Growth and natural mortality were sex-specific and time invariant. Sample sizes were iteratively re-weighted using the Francis (2011) method based on equation TA1.8. The value of σ_R was set at 0.7. Additional variances were estimated for the Triennial and NWFSC shelf-slope survey indices. For the base model brought to the STAR, steepness was fixed at 0.5, and natural mortality was set at 0.054 year^{-1} for both females and males.

Selectivities for each fishery and survey, which were length-dependent, time invariant, and common to both females and males, were modeled using the double normal pattern, except in the case of the Pacific ocean perch survey, for which logistic selectivity was assumed. Subsequent exploration of the assessment model determined that, although domed shape selectivity provided the best fit to the data for the fishery fleet and Triennial survey, selectivities of the at-sea hake fishery and all other surveys were best described using asymptotic selectivity curves. Fishery retention was time-blocked to reflect the effects of the various management changes that the fishery had experienced. Catchabilities were estimated analytically. Recruitment deviations from 1900 were estimated. Little or no bias adjustment for recruitment was applied from 1900 to 1974, with an upward and downward ramping of bias adjustment in the main period of recruitment deviations from 1974 to 2014, and no bias adjustment for 2015 onward.

The final agreed base model differed slightly from the model which had originally been proposed in the draft assessment report. Triennial survey indices and associated compositions had been removed, fishery selectivity and retention time blocks had been modified, steepness had been set to the mean from the steepness meta-analysis, i.e., 0.72, and the model had been re-tuned.

Pacific ocean perch ToR 2. Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.

The approaches used to analyse and collate the data used for the 2017 assessment of Pacific ocean perch were described in the draft assessment report and discussed during the review. It was noted during these discussions that Pacific ocean perch discarded by the shrimp fishery are not included in the discards considered in the assessment model. Although the biomass of such discards is trivial, the numbers discarded might be non-trivial. The annual numbers of such discarded Pacific Ocean perch were estimated by the STAT during the review. Following discussion, it was concluded by the STAR Panel that inclusion of these data within the 2017 assessment model was not necessary.

The assumption of the CAAL data for the NWFSC shelf-slope survey is that, within an individual sample, the ages at a given length are representative of the distribution of ages at that length within the population. Such an assumption is inconsistent with the understanding that the spatial distribution of the Pacific ocean perch is age-dependent. Further investigation is needed of age-dependent movement and distributions of Pacific ocean perch, and of the implications of such age dependence on growth estimates and on the way in which the CAAL data are treated in SS3.

Application of the spatio-temporal modelling approach in VAST provided an interesting alternative approach to estimating indices of abundance than the traditional design-based approaches. Apparently samples of composition data are still expanded to an overall composition using a traditional approach. This results in a slight mis-match between the indices and compositions, which should be investigated.

Length compositions for discards from the commercial fishery in 1986 included not only data for both sexes but also data for some unsexed fish. This would be an issue when dealing with sex ratio. The length composition data should be input either as sexed or unsexed compositions, not a mixture of the two.

Pacific ocean perch ToR 3. Evaluate model assumptions, estimates, and major sources of uncertainty.

In the initial presentation of the Pacific ocean perch assessment to the STAR, the STAT advised of several corrections to data that would be made for the post-STAR model, and that a slight change had been made to the SS3 data file in order that the surveys are treated as fleets, thereby removing the catches, which otherwise do not get removed. Essentially no difference to spawning output was apparent when the model was run with the modified data.

In contrast to the approach used for lingcod, where the population was assumed to be at an unexploited equilibrium at the start of the modeled period, the model for Pacific ocean perch was started and run with zero catch from 2000 such that uncertainty in the ratio of spawning output to unexploited spawning output at the start of fishing in 2018 was more realistically represented.

The extraordinarily high recruitment predicted for 2008 was initially a source of concern to the Panel. Based on model outputs, however, it appears certain that recruitment in 2008 was very strong (even in log space), but there is large uncertainty regarding the actual number of recruits. There are early indications of a strong recruitment also in 2013.

The rapid increase in spawning output at the end of the series was also of concern and was investigated by the Panel. A sensitivity run that set the recruitment deviations for 2008 and 2013 to zero demonstrated that the large upturn in spawning output at the end of the series is due to these strong levels of annual recruitment. Results of model runs in which data components were successively introduced demonstrated that the upturn of spawning output at the end of series is partly due to the addition of the new ages and lengths to the data used in 2011, not the influence of the new catches or surveys in this period. The key difference between the overall pattern of the trajectory of spawning output from the 2011 assessment and that of the 2017 assessment stems from the changes that were made to *h* and *M*.

CAALs for the NWFSC shelf-slope survey were initially input to the model to allow estimation of growth. In a sensitivity run when growth was then input and fixed, and the CAALs replaced by the marginal age composition for this survey (with likelihoods of both the length and age compositions halved, to account for 'double counting', then re-tuned), other parameters were affected by the change. That is, because of the way in which SS3 balances the log-likelihoods of the different data sets, the influence of the CAALs extends beyond the estimates of the growth parameters to the other parameters of the model, thereby affecting estimates of population age structure, trajectories of spawning output and ultimately status determination.

Pacific ocean perch ToR 4. Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.

Improvements to the proposed base model presented in the draft assessment report were identified during the course of the review following exploration of the influences of the various data components.

A key uncertainty was associated with the conflict between the information in the NWFSC shelf-slope survey indices and the information within the Triennial survey indices and associated compositions. The latter data scaled spawning output to a much lower level than the former, with greater depletion. The NWFSC survey was considered by the STAT to be the more reliable of the two indices, i.e., the survey most representative of the Pacific ocean perch population. Accordingly, it was retained within the final base model and the Triennial survey indices and associated compositions removed. It was recommended that a sensitivity run be included in the final assessment report to recognise the alternate state of nature represented by inclusion of the Triennial survey data.

Modification of the selectivity and retention time blocks for the fishery was also recommended to provide a better representation of the size composition of the catches in recent years. It was noted that relatively few additional parameters would be required when adding the additional time blocks.

Pacific ocean perch ToR 5. Determine whether the science reviewed is considered to be the best scientific information available.

The final agreed base model for the Pacific ocean perch population was accepted by the STAR as well structured, had been thoroughly investigated by the STAT, and is considered likely to provide the best scientific information currently available for the formulation of management advice.

Pacific ocean perch ToR 6. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.

Short term

An investigation should be undertaken to determine whether, for consistency, the approach used in surveys for expansion of composition data should be modified from the traditional design-based approach when indices of abundance are calculated using the VAST procedure. Further exploration of the implications of use of the VAST approach rather than the design-based approach when constructing indices of abundance is also warranted.

A study of existing survey and fisheries data should be undertaken to assess whether, for Pacific ocean perch of either sex or combined sexes (for samples where it is not possible to cut fish to determine the sexes of the fish), conditional age distributions at different lengths vary spatially or temporally, thereby exploring the assumption that, in conditional age-at-length data, ages of sampled fish at a given length are random samples of the distribution of ages at that length in the population. If non-random, construction of marginal length compositions should account appropriately for the non-uniform sampling of age by length and the implications for use of CAAL data in the assessment model and in constructing growth curves should be considered.

Long term

Results of genetic studies are consistent with the possibility of transboundary connectivity between population in waters off US west coast and Canada. The STAT has recognised this by including a sensitivity run to explore the effect of inclusion of the Canadian fishery and survey data that are available. It has acknowledged that this should be an ongoing area of research and collaboration between the US and Canada. This view is endorsed by the STAR Panel. Data should be examined to determine the extents of correlations among indices of abundance, and among recruitment deviations, for Pacific ocean perch from the US west coast, Canada, and Alaska.

The conflict between the information contained in the NWFSC shelf-slope survey indices and that contained in the Triennial survey indices and compositions requires further investigation. Is the inconsistency due to survey design or implementation, or is it due to the spatial distribution and catchability of the individuals of different sizes, ages, and sexes? Does the issue extend to other species?

Pacific ocean perch ToR 7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

Estimates of spawning output produced using the 2017 base model proposed in the draft assessment report suggested that depletion of the Pacific ocean perch population was not as great as that estimated with the 2011 assessment model. An exploration of the incremental consequences of updating catch and discard data to 2017, updating the indices, updating the length compositions, and updating the age compositions demonstrated that the length and age data were responsible for some of the increase in the most recent years, presumably as a result of the strong recruitment that was evident in these data, but that much of the improvement in estimates of spawning output was the result of the changes to natural mortality and steepness.

Although small in biomass, considerable numbers of tiny Pacific ocean perch are potentially removed from the population by the shrimp fishery. The STAT estimated the numbers removed annually by the shrimp fishery between 2004 and 2015 as ranging from 1,300-310,000 fish. It was not considered necessary to include these removals in the assessment model.

Clarification of the approach employed in expansion from NWFSC samples to marginal compositions was sought. Expansions are based on numbers per hectare.

The STAT advised that the data source likely to provide the best representation of the dynamics of the Pacific ocean perch population was the NWFSC survey. The STAT was asked to run the assessment model with the NWFSC index, as estimated using VAST, but with no other indices of abundance, and without estimation of additional SD. The run was to be repeated using the design-based estimates of the NWFSC index. This run would remove possible confusing signals from other indices, and allow identification of changes to the spawning output trajectory produced by using the VAST approach rather than the design-based approach when estimating the indices of abundance for the NWFSC survey. It was concluded that use of the VAST rather than the design-based approach of analysis of survey data had made little difference to assessment results, and that, when run with only the NWFSC index, spawning output remained above the management target throughout the series, i.e. was scaled to a higher level than the estimates produced using the proposed base model.

A subsequent run in which only the NWFSC and Triennial surveys, but no other indices of abundance were included, demonstrated that the lower scale of the spawning output of the base model was probably due to the influence of the Triennial survey, which extended the period of survey coverage.

The SS3 model was run as an Age Structured Production Model (ASPM) where all estimated parameters from the candidate base model were fixed (specified and not estimated) at the base case estimates, with the exception of global scaling parameters $\ln(R_0)$ and q, which were estimated. Recruitment deviations were not estimated and recruitment was taken from the spawner-recruit relationship. Composition data were turned off in the model and only key indices of abundance contributed to the total likelihood. No extra SD was estimated in any of these runs. The resultant spawning output was predicted to remain at a much higher level, relative to the management target, throughout the modelled period, despite the inclusion of both the NWFSC and Triennial surveys. It was concluded that the composition data, rather than the survey indices, were the key determinants of the scaling relative to the management target.

The effect of the CAAL data on estimates of other parameters, even when growth parameters were fixed was discussed. The likelihood associated with the CAAL data influenced not only the parameter estimates for growth equations, but also the parameter estimates associated with other aspects of the model. If growth is fixed, but the likelihood associated with the CAALs continues to contribute to the total likelihood, the CAALs will continue to influence estimates of other parameters. The model was run with growth parameters fixed at those of the base model but without including the contribution of the CAAL to the total likelihood. Spawning output was scaled lower than that predicted by the base model, and depletion in 2017 was greater than that predicted by the base model. It was concluded that inclusion of the CAAL data in the model to estimate growth parameters would influence estimates of spawning output. The possibility that un-modeled age-dependent movement and spatial distribution are biasing the CAAL data was discussed.

As expected, a model run using a value of M corresponding to a maximum age of 120 years produced spawning output estimates that were lower, with a greater depletion, than those predicted by the base model.

The model was run with the 2008 recruitment deviation set to 0, and details of likelihood components were considered. The data source providing the information responsible for the extraordinarily large 2008 year class estimate was identified as the length composition data from the NWFSC shelf-slope survey.

The STAT was asked to modify the data in accordance with the necessary changes and corrections that had been identified prior to the review and to run the model without the fishery CPUE index, for which details of analysis had been lost and which could not be reproduced. The modifications were made, together with a correction to the sample size for the NWFSC CAAL data, which had been identified during the review. These changes, and removal of the CPUE index, made little difference to spawning output. The STAR Panel agreed with the STAT that it was appropriate that the result of this run should become the foundation of the new base model for Pacific ocean perch.

Using this revised model, the effect of splitting the Triennial survey data was explored to determine whether there was evidence in the Triennial data of a change in q. The estimate of q for the earlier Triennial data was greater than that for the later data. With no evidence of a change in survey design or implementation, however, there was no basis for splitting the series.

To determine the cause for the lack of fit to the early Triennial survey, the revised model was run with only the Triennial survey and fishery composition data, with growth fixed. If time permitted, the STAT was requested to attempt to estimate steepness, using the prior, to determine whether, if data conflicts were reduced, steepness was estimable. The requested run resulted in far lower estimates of spawning output, and far greater depletion than the revised base model. Estimates of steepness were reduced from the fixed value of 0.5 in the revised base model to ~0.3.

Again using the revised base model, the use of dome-shaped age selection for the NWFSC shelf-slope survey was explored. The purpose of this run was to explore the consequences of violations of the CAAL assumption that ages are random at length in the survey. The resulting age-based selectivity curves were asymptotic, with approximately uniform selectivity over all ages. In discussion, however, it was questioned whether sufficient information on population age composition was likely to be available from other data sources to allow determination of the form of the age selectivity curve for the NWFSC CAAL data.

The possibility that the lack of fit of the Triennial survey data might be the result of time-varying selectivity was discussed. The STAT was requested to investigate whether varying the form of selectivity in the revised base model might improve model fit. Results of exploration suggested that there may have been a shift of selectivity towards smaller fish and change in retention width in more recent years. The STAT proposed that a new time-block from 2000 to 2016 be included in the model to allow for this change in selectivity and retention.

The STAT considered a request from the STAR Panel that it should work towards a revised base and, noting that the original assumption that h=0.5 might be arbitrary, provide a steepness sensitivity analysis for a new preferred base. In response, the STAT proposed that the new preferred base should be based on the current revised base model, with the time-blocks for selectivity and retention in recent years that had been identified in the immediately preceding exploration, but with removal of the Triennial survey and its composition data. It advised that, as in previous models, the new preferred base model should continue to use the CAAL from NWFSC shelf slope, and proposed that steepness should be set to the value of the current prior for rockfish, i.e., h=0.72.

To assess the implications of removing the Triennial survey in favour of the NWFSC survey in the new preferred base model, the STAR Panel requested a model run with only the Triennial survey (but no added variance) and associated composition data, including fishery length compositions. It was requested that the run should assume h = 0.4 and alternatively h = 0.5 and 0.72, re-weighting with h= 0.5 before running the other steepness scenarios. The predicted decline in spawning output was much more severe under all three steepness scenarios than that predicted using the base model, with the 2016 output falling below the minimum stock size threshold in all three cases. The decision to remove the Triennial survey data from the preferred new base model will result in more optimistic estimates of stock status and depletion.

The STAT provided a plot of fishery selectivity and retention for the 2011 assessment with the time block ending at 2010, comparing this with the selectivity patterns estimated for the proposed base model for 1918-1999 and 2000-2016 and the retention for 2010. Selectivity of larger fish was much lower following 2000 than in the selectivity patterns for the 2011 assessment and in the period 1918-1999. Retention in 2010 in the 2017 assessment was also lower than in the earlier assessment. Thus, the 2017 assessment predicts that fewer large fish are being caught and that there is greater retention of the smaller fish.

The STAT explored a potential base model, which they had proposed, i.e., a model from which Triennial survey and associated compositions had been removed, fishery selectivity and retention time blocks had been modified, steepness had been set to the mean from the steepness meta-analysis (0.72). Full diagnostic output was produced and likelihood profiles were produced for fixed values of female M, h, and $\ln(R_0)$. Following consideration of the diagnostic output, the STAT advised that this model should be accepted as the final base model for the 2017 assessment of Pacific ocean perch. It was agreed that a decision table using values of steepness consistent with the 12.5^{th} and 87.5^{th} percentiles of the asymptotic distribution of spawning output from the base model assessment would be produced by the STAT.

To inform the development of the decision table, the STAT was requested to explore quantiles on steepness from Thorson's meta-analysis of h for the major axis of uncertainty in the decision table. The low and high values of h corresponding to the 12.5^{th} and 87.5^{th} percentiles were estimated to be 0.509 and 0.894, respectively. The STAT was also requested to explore quantiles on M for the major axis of uncertainty. For this, values of M corresponding to the 12.5^{th} and 87.5^{th} percentiles were estimated to be 0.0325 and 0.089, respectively.

Time did not allow the STAR Panel to examine the results of this final request. No issues were detected when examining the section describing the production of the decision table within the draft report, which was subsequently produced by the STAT.

5. Conclusions and recommendations

The NMFS review process allows a detailed and open appraisal of the data that enter an assessment and the assessment models employed to analyse those data and produce the advice required for consideration by fisheries managers. Through incremental review and response to constructive criticism, the assessments are continually improved. The assessment documents that are produced for each assessment provide a valuable and detailed record that is invaluable for those who are seeking to understand the basis for management advice and decisions, and why research funding is directed towards particular studies. Although the review process has been refined over time, and is now very efficient, there is still opportunity for a few small improvements, such as those I identify below.

Although draft assessment reports can only be disseminated shortly before the scheduled NMFS review meeting, many of the background documents are available well before that meeting, e.g., published scientific papers, reports from previous assessments. Given the volume of documents that must be examined prior to a review, it would be useful if those papers that are already ready for distribution could be made available on an FTP site when panelists are first appointed. This would allow panelists to familiarise themselves with the background material prior to receiving the draft assessment reports when these eventually become available.

Use was made in the review of a laser pointer when drawing attention to particular features of a figure or table. It would assist remote webinar attendees if the cursor on the computer screen was used to identify those features, rather than the laser pointer.

6. Reference list

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68:1124–1138.

Gunderson, D. R., and Sample, T. M. 1980. Distribution and abundance of rockfish off Washington, Oregon and California during 1977. Northwest; Alaska Fisheries Center, National Marine Fisheries Service. Available from http://spo.nmfs.noaa.gov/mfr423-4/mfr423-42.pdf.

Hamel, O. S. 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES Journal of Marine Science: Journal du Conseil 72(1): 62{69. doi: 10.1093/icesjms/fsu131.

Stewart, I. J., and Hamel, O. S. 2014. Bootstrapping of sample sizes for length- or age composition data used in stock assessments. Canadian Journal of Fisheries and Aquatic Sciences 71(4): 581{588. doi: 10.1139/cjfas-2013-0289.

Then, A. Y., Hoenig, J. M., Hall, N. G., and Hewitt, D. A. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science 72(1): 82{92. doi: 10.1093/icesjms/fsu136.

Appendix 1: Bibliography of materials provided

Background – All STAR Panels

Anon. 2017. Report of the Groundfish Historical Catch Reconstruction Workshop. Pacific Fishery Management Council. Portland, OR. 1-3 November 2016.

Anon. 2017. Report of the Groundfish Productivity Workshop of the Pacific Fishery Management Council's Scientific and Statistical Committee. NOAA Fisheries, Alaska Fisheries Science Center, Seattle, Washington. December 6-8, 2016.

Anon. 2017. Scientific and Statistical Committee Report on Reports and Recommendations from Groundfish Science Workshops and Methodology Reviews.

Anon. 2017. Summary Minutes. 2017 Groundfish Pre-Assessment Workshop. March 21-22, 2017. Pacific Fishery Management Council.

Anon. No date. Accepted Practices Guidelines for Groundfish Stock Assessments.

Anon. No date. Summary table of Triennial Shelf Survey and NWFSC Shelf/Slope Survey. This table was completed with input from Victor Simon and Aimee Keller (NWFSC) and Mark Wilkins (AFSC).

Dorn. M. 2016. Design considerations for PFMC groundfish stock assessments and reference points. Powerpoint presentation. Groundfish Productivity Workshop. Seattle, Washington.

Hamel, O. S. 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. *ICES Journal of Marine Science*, **72**: 62–69.

Hamel, O. S. 2017. An update to developing a prior for the natural mortality rate for fishes for use in the 2017 assessment cycle for the U.S. West Coast. Unpublished document.

Karnowski, M., Gertseva, V., and Stephens, A. 2014. Historical Reconstruction of Oregon's Commercial Fisheries Landings. Fish Division, Oregon Department of Fish and Wildlife. Information Reports Number 2014-02. 52 pp.

NOAA. No date. Information Quality Guidelines.

Pacific Fishery Management Council. 2016. Terms of Reference for the Groundfish and Coastal Pelagic Species Stock Assessment Review Process for 2017-2018.

Thorson, J. T. 2017. Estimating a Bayesian prior for steepness in Pacific rockfishes (*Sebastes* spp.) off the U.S. West Coast for the 2017 assessment cycle. Presented to SSC Assessment Methods Review. Jan. 11, 2017.

Thorson, J. T., and Barnett, L. A. K. (2017) Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. – ICES Journal of Marine Science, doi:10.1093/icesjms/fsw193.

Thorson, J. T., Shelton, A. O., Ward, E. J., and Skaug, H. J. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. – ICES Journal of Marine Science, doi: 10.1093/icesjms/fsu243.

Thorson, J. T., Dorn, M. W., and Hamel, O. S. No date. Steepness for West Coast rockfishes: Results from a twelve-year 1 experiment in iterative regional meta-analysis. Draft: For Review During PFMC Star Panel Reviews In 2017.

Background – STAR Panel 1

Anon., No date. Summary of California Recreational Management Measures Affecting Lingcod.

Anon. No date. [Federal] Regulations [affecting Lingcod] by Multiple Parameters.

Anon. No date. Excel workbook, Lingcod_SupportingMaterials_WA_Lingcod-RecRegsArea1-4.xlsx. [Annual regulations for recreational fishing for Lingcod].

Chen, Y. 2011. CIE Independent Peer Review Report on Pacific Ocean Perch and Petrale Sole STAR Review. Report to CIE.

Hamel, O. S., Sethi, S. A., and Wadsworth, T. F. 2009. Status and Future Prospects for Lingcod in Waters off Washington, Oregon, and California as Assessed in 2009. Lingcod Assessment 2009 – Final – SAFE Version.

Hamel, O. S. 2009. Status and Future Prospects for the Pacific Ocean Perch Resource in Waters off Washington and Oregon as Assessed in 2009. Pacific Ocean Perch Assessment 2009 – Final – SAFE Version.

Hamel, O. S., and Ono, K. 2011. Stock Assessment of Pacific Ocean Perch in Waters off of the U.S. West Coast in 2011.

Jagielo, T. H. and Wallace, F. R. 2005. Assessment of Lingcod (*Ophiodon elongatus*) for the Pacific Fishery Management Council in 2005.

Maguire, J.-J. 2009. Report on the cabezon and lingcod Stock Assessment Review (STAR) Panel, July 27 - 31, 2009, Seattle, WA. Report to CIE.

Smith, S. 2009. Report for the Center of Independent Experts on the Stock Assessment Review (STAR) Panel for Cabezon and Lingcod (July 27 to 31, 2009). Report to CIE.

SSC. 2009. Scientific and Statistical Committee Report on Stock Assessments for 2011-2012 Groundfish Fisheries. Supplemental SSC Report, June 2009.

STAR Panel. 2005. STAR Panel Report. Lingcod. Northwest Fisheries Science Center, Seattle, Washington, USA. 15-19 August 2005.

STAR Panel. 2005. Lingcod. STAR Panel Report. Alaska Fisheries Science Center, Seattle, Washington, USA. September 26-30, 2005.

STAR Panel. 2009. Lingcod. STAR Panel Report. Deca Hotel, Seattle, Washington. July 27-30, 2009.

STAR Panel. Pacific Ocean Perch. Stock Assessment Review (STAR) Panel Report. Hotel Deca, Seattle, Washington. 20-24 June 2011.

Stokes, K. 2011. Report on the Stock Assessment Review (STAR) Panel for Pacific Ocean Perch and Petrale Sole. Report to CIE.

Whitman, A. 2017. Summary: Sport Regulation Changes relevant to Lingcod, and other items that may be relevant to CPUE development.

Draft Lingcod Assessment

4_North_v7-hess-ToSTAR. Folder containing input files for SS3 analysis of northern stock of lingcod and associated output tables and figures.

4_North_w2.1-hess-AltToSTAR. Folder containing alternate input files.

5_South_e_base-hess-ToSTAR. Folder containing input files for SS3 analysis of southern stock of lingcod and associated output tables and figures.

5-South_f5-hess-AltToSTAR. Folder containing alternative input files.

Haltuch, M. A., Wallace, J., Akselrud, C. A., Nowlis, J., Barnett, L. A. K., Valero, J. L., Tsou, T.-S., and Lam, L. 2017. Draft 2017 Lingcod Stock Assessment, with associated figures, tables, and Excel file containing parameters.

Future Research Recommendations for the Lingcod Assessment (developed during STAR).

Draft Pacific Ocean Perch Assessment

CSV files, POP_natage_f.csv and POP_natage_m.csv, containing number-at-age data for Pacific Ocean Perch.

SS3 input files for Pacific Ocean Perch stock assessment.

Recommendations for Future Research for the POP Assessment (developed during STAR).

Wetzel, C. R. and Cronin-Fine, L. 2017. Status of Pacific ocean perch (*Sebastes alutus*) along the US west coast in 2017.

Stock Synthesis Documentation

Methot Jr, R. D. and Wetzel, C. R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, **142**: 86– 99.

Methot Jr, R. D. and Wetzel, C. R. 2013. Appendix A: Technical Description of the Stock Synthesis assessment program. Supplementary data. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, **142**: 86–99.

Methot, R. D., A'mar, T., Wetzel, C., and Taylor, I. 2017. Stock Synthesis. User Manual. Version 3.30.04. June 1, 2017. NOAA Fisheries, Seattle, WA USA.

Appendix 2: Copy of the CIE Statement of Work

External Independent Peer Review by the Center for Independent Experts

Stock Assessment Review (STAR) Panel 1

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).

Further information on the CIE program may be obtained from www.ciereviews.org.

Project Description

The National Marine Fisheries Service and the Pacific Fishery Management Council will hold stock assessment review (STAR) panels in 2017 to evaluate and review benchmark assessments of Pacific coast groundfish stocks. The goals and objectives of the groundfish STAR process are to:

1) ensure that stock assessments represent the best available scientific information and facilitate the use of this information by the Council to adopt OFLs, ABCs, ACLs, (HGs), and ACTs;
2) meet the mandates of the Magnuson-Stevens Fisheries Conservation and Management Act (MSA) and other legal requirements;
3) follow a detailed calendar and fulfill explicit responsibilities for all participants to produce required reports and outcomes;
4) provide an independent external review of stock assessments;
5) increase understanding and acceptance of stock assessments and peer reviews by all members of the Council family;
6) identify research needed to improve assessments, reviews, and fishery management in the future; and
7) use assessment and review resources effectively and efficiently.

Benchmark stock assessments will be conducted and reviewed for lingcod and Pacific ocean perch. Lingcod has been an important groundfish target species along the west coast of the United States, ranking in the top-6 of importance for commercial, recreational, and tribal fisheries. This will be the first assessment for lingcod since a benchmark assessment was completed in 2009. During the last benchmark assessment, the sensitivity results showed high uncertainty in age data as well as in the status of the population in southern California. A substantial effort is currently underway to age lingcod to reduce uncertainties relating to age and growth data and improve recruitment estimates.

Pacific ocean perch has been managed under a rebuilding plan for over a decade and, while not expected to be rebuilt for several more decades (2051), was identified as a strong candidate for assessment during the Pacific coast groundfish regional stock assessment prioritization process, which was based on the national stock assessment prioritization framework (http://www.st.nmfs.noaa.gov/Assets/stock/documents/PrioritizingFishStockAssessments_FinalWeb.pdf. Pacific ocean perch was assessed as a benchmark assessment in 2011 and a catch-only rebuilding projection update in 2015 to monitor the rebuilding progress and provide updated scientific-based advice for management. This stock has recently become more constraining on elements of the Pacific hake fishery.

Assessments for these two stocks will provide the basis for the management of the groundfish fisheries off the West Coast of the U.S. including providing scientific basis for setting OFLs and ABCs as mandated by the Magnuson-Stevens Act. The technical review will take place during a formal, public, multiple-day meeting of fishery stock assessment experts. Participation of external, independent reviewer is an essential part of the review process. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers

NMFS requires two CIE reviewers to participate in this stock assessment review panel. One CIE reviewer shall conduct an impartial and independent peer review of the two assessments described above and in accordance with the SoW and ToRs herein. Additionally, a second "consistent" CIE reviewer will participate in all STAR panels held in 2017 and the SOW and ToRs for the "consistent" CIE reviewer are included in a separate SoW (See **Attachment A**).

Both CIE reviewers shall be active and engaged participants throughout panel discussions and able to voice concerns, suggestions, and improvements while respectfully interacting with other review panel members, advisors, and stock assessment technical teams. The CIE reviewers shall have excellent communication skills in addition to working knowledge and recent experience in fish population dynamics, with experience in the integrated analysis modeling approach, using age-and size-structured models, use of *Markov Chain Monte Carlo (*MCMC) to develop confidence intervals, and use of Generalized Linear Models in stock assessment models.

Statement of Tasks

The CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Pre-review Background Documents: At least two weeks before the peer review, the contractor will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewer shall read all documents in preparation for the peer review.

Documents to be provided to the CIE reviewers prior to the STAR Panel meeting include:

- The current draft stock assessment reports;
- The Pacific Fishery Management Council's Scientific and Statistical Committee's Terms of Reference for Stock Assessments and STAR Panel Reviews;
- Stock Synthesis (SS) Documentation
- Additional supporting documents as available (including previous stock assessments and STAR panel reports).
- An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer).

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein.

Contract Deliverables - Independent CIE Peer Review Reports: The CIE reviewers shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**.

Other Tasks – Contribution to Summary Report: The CIE reviewers may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. The CIE reviewers are not required to reach a consensus, and should provide a brief summary of each reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Timeline for CIE Reviewers
The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided in advance of the peer review.
2) Participate during the STAR Panel 1 review meeting in **scheduled in Seattle, Washington during the dates of June 26-30, 2017** as specified herein, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than **July 14, 2017**, each CIE reviewer shall submit their draft independent peer review report to the contractor. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each ToR in **Annex 2**.

Foreign National Security Clearance
When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website:

http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
The CIE reviewers shall conduct an independent peer review during the panel review meeting scheduled in **Seattle, Washington during the dates of June 26-30, 2017**.

Period of Performance
The period of performance shall be from the time of the award through August 18, 2017. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables
The contractor shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

May 22, 2017	Contractor selects and confirms reviewers
June 12, 2017	Contractor provides the pre-review documents to the reviewers
June 26-30, 2017	Each reviewer participates and conducts an independent peer review during the panel review meeting
July 14, 2017	Contractor receives draft reports
July 31, 2017	Contractor submits final reports to the Government

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards: (1) The reports shall be completed in accordance with the required formatting and content in **Annex 1**; (2) The reports shall address each ToR as specified **Annex 2**; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $8,200.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contacts:
Stacey Miller, NMFS Project Contact
National Marine Fisheries Service,
2032 SE OSU Drive
Newport, OR 97365
Stacey.Miller@noaa.gov

Phone: 541-867-0535

Jim Hastie
National Marine Fisheries Service,
2725 Montlake Blvd. E,
Seattle WA 98112
Jim.Hastie@noaa.gov
Phone: 206-860-3412

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

Stock Assessment Review (STAR) Panel 1

1. Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.
2. Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.
3. Evaluate model assumptions, estimates, and major sources of uncertainty.
4. Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.
5. Determine whether the science reviewed is considered to be the best scientific information available.
6. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.
7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

Annex 3: Agenda for STAR Panel 1 Review Meeting

June 26-30, 2017

Proposed Agenda
Stock Assessment Review (STAR) Panel 1
for Lingcod and Pacific Ocean Perch

NOAA Fisheries, Northwest Fisheries Science Center
2725 Montlake Blvd. East
Seattle, WA 98112

June 26-June 30, 2017

This meeting is open to the public and public comments from attendees will be accepted at the discretion of the meeting chair. Agenda times are approximate and are subject to change.

Monday, June 26

8:30 a.m.	Welcome and Introductions
9:00 a.m.	Review the Draft Agenda and Discuss Meeting Format (Chair) - Review the Terms of Reference (TOR) for assessments, Accepted Practices Guidelines, and STAR panel responsibilities - Assign reporting duties - Agree on time and method for accepting public comments
9:30 a.m.	Presentation of the Lingcod Assessment - Overview of data and modeling
12:30 p.m.	Lunch (On Your Own)
2:00 p.m.	Q&A session with Lingcod Stock Assessment Team (STAT) STAR Panel discussion - Panel develops written request for additional model runs / analyses
4:00 p.m.	Begin Presentation of the POP Assessment (if time allows) - Overview of data and modeling
5:30 p.m.	Adjourn for day

Tuesday, June 27

8:30 a.m.	Continue and Complete Presentation of the POP Assessment - Overview of data and modeling
10:00 a.m.	Q&A Session with the POP STAT Panel Discussion - Panel develops written request for additional model runs / analyses
12:00 p.m.	Lunch (On Your Own)
1:30 p.m.	Begin Drafting the STAR Panel Report
4:00 p.m.	Begin Presentation of the First Set of Requested Model Runs by the Lingcod STAT
5:30 p.m.	Adjourn for day

Wednesday, June 28

8:30 a.m.	Presentation of the First Set of Requested Model Runs by the Lingcod STAT - Q&A session with the Lingcod STAT & Panel discussion - Panel develops request for second round of model runs / analyses for the Lingcod STAT
10:30 a.m.	Begin Presentation of the First Set of Model Runs by the POP STAT
12:00 p.m.	Lunch
1:30 p.m.	Continue Presentation of the First Set of Model Runs by the POP STAT - Q&A session with the POP STAT & panel discussion - Panel develops request for second round of model runs / analyses for the POP STAT.
3:30 p.m.	Continue Drafting STAR Panel Report
5:30 p.m.	Adjourn for day

Thursday, June 29

8:30 a.m.	Presentation of the Second Set of Model Runs by the Lingcod STAT - Q&A session with the Lingcod STAT & panel discussion - Agreement of the preferred model and model runs for the decision table - Panel continues drafting the STAR report.
12:00 p.m.	Lunch (On Your Own)
1:30 p.m.	Presentation of the Second Set of Model Runs by the POP STAT - Q&A session with the POP STAT & panel discussion - Agreement of the preferred model and model runs for the decision table - Panel continues drafting the STAR report.
5:00 p.m.	Continue Panel Discussion or Drafting of the STAR Panel Report
5:30 p.m.	Adjourn for day

Friday, June 30

8:30 a.m.	Consideration of Remaining Issues - Review decision tables for assessments
10:00 a.m.	Panel Report Drafting Session
12:00 p.m.	Lunch (on your own)
1:30 p.m.	Review First Draft of the STAR Panel Report
4:00 p.m.	Panel Agrees to Process for Completing the Final STAR Report for Council's September Meeting Briefing Book (Requested by August 14th)
5:30 p.m.	Review Panel Adjourns

Appendix 3: Participant List

The list below was provided in the background documentation for STAR Panel 1.

Participants
Stock Assessment Review Panel for
Lingcod and Pacific Ocean Perch

NOAA Fisheries Northwest Fisheries Science Center
2725 Montlake Blvd. East
Seattle, Washington 98112

June 26-30, 2017

Technical Reviewers
David Sampson, Scientific and Statistical Committee (SSC), Panel Chair
Panayiota (Yiota) Apostolaki, Center for Independent Experts (CIE)
Norm Hall, Center for Independent Experts (CIE)
Kevin Piner, Scientific and Statistical Committee (SSC)

Panel Advisors
John DeVore, Pacific Fishery Management Council (PFMC), Staff Officer, PFMC
Lynne Mattes, Groundfish Management Team (GMT)
Louie Zimm, PFMC Groundfish Advisory Subpanel (GAP)

Stock Assessment (STAT) Teams

Lingcod STAT
Melissa Haltuch, Northwest Fisheries Science Center
John Wallace, Northwest Fisheries Science Center
Caitlin Allen Akselrud, University of Washington
Josh Nowlis, Northwest Fisheries Science Center
Lewis Barnett, Northwest Fisheries Science Center and University of Washington
Juan Valero, Center for the Advancement of Population Assessment Methodology
Tien-Shui Tsou, Washington Department of Fish and Wildlife
Laurel Lam, Moss Landing Marine Laboratories

Pacific Ocean Perch STAT
Chantel Wetzel, Northwest Fisheries Science Center
Lee Cronin-Fine, University of Washington

Numerous other individuals, who, at various times, were present or monitoring the proceedings, participated and offered valuable comment during the review meeting.

Appendix 4. Estimation of confidence limits from likelihood profile

Haddon (2011) advises that "the expected log-likelihoods for the confidence limits for a single parameter θ, assuming all other parameters remain at their optimum, are given by

$$LL(\theta) = LL(\theta)_{\mathrm{Max}} - \frac{\chi^2_{1,1-\alpha}}{2}$$

where $\chi^2_{1,1-\alpha}$ is the $(1-\alpha)$th quantile of the χ^2 distribution with 1 degree of freedom".

95% confidence limits

For 95% confidence limits, $\alpha = 0.05$, $1 - \alpha = 0.95$, and $\chi^2_{1,1-\alpha} = 3.84$. The 95% confidence limits for θ are then determined as the values of θ at which the likelihood profile for θ intersects the line

$$LL(\theta) = LL(\theta)_{\mathrm{Max}} - \frac{3.84}{2} = LL(\theta)_{\mathrm{Max}} - 1.92.$$

If use of negative log-likelihoods (where NLL = - LL) is preferred, the confidence limits would be determined as the values for θ at the intersection of the profile curve for the negative log-likelihood and the line

$$NLL(\theta) = NLL(\theta)_{\mathrm{Min}} + \frac{3.84}{2} = NLL(\theta)_{\mathrm{Min}} + 1.92.$$

Values of parameter at 12.5th and 87.5th percentiles for alternative states of nature

The values of the 12.5th and 87.5th percentiles of the distribution of the values of the parameter of interest represent the parameters that may be used to generate the alternative states of nature that bracket the state of nature associated with the maximum likelihood estimate of the parameter. These three states of nature are used to produce the decision table containing the information required for management.

The 12.5th and 87.5th percentiles of this distribution represent the confidence limits when $\alpha = 0.25$.

For 75% confidence limits, $\alpha = 0.25$, $1 - \alpha = 0.75$, and $\chi^2_{1,1-\alpha} = 1.32$. The 75% confidence limits for θ are then determined as the values of θ at which the likelihood profile for θ intersects the line

$$LL(\theta) = LL(\theta)_{\mathrm{Max}} - \frac{1.32}{2} = LL(\theta)_{\mathrm{Max}} - 0.66.$$

If use of negative log-likelihoods is preferred, the confidence limits would be determined as the values for θ at the intersection of the profile curve for the negative log-likelihood and the line

$$NLL(\theta) = NLL(\theta)_{\mathrm{Min}} + \frac{1.32}{2} = NLL(\theta)_{\mathrm{Min}} + 0.66.$$

Reference list

Haddon, M. 2011. Modelling and Quantitative Methods in Fisheries. Second Edition. Chapman and Hall/CRC.